PCT/AU2004/000768

Australian Government

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patent Office
Canberra

I, JULIE BILLINGSLEY, TEAM LEADER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2003902907 for a patent by PROTEOME SYSTEMS INTELLECTUAL PROPERTY PTY LTD as filed on 11 June 2003.

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

WITNESS my hand this
Twenty-first day of June 2004

J. Billingsley

JULIE BILLINGSLEY
TEAM LEADER EXAMINATION
SUPPORT AND SALES

AUSTRALIA

Patents Act 1990

Proteome Systems Intellectual Property Pty Ltd

PROVISIONAL SPECIFICATION

Invention Title:

Characterisation of glycans

The invention is described in the following statement:

Field of the Invention

This invention relates to a method of characterising glycans and their derivatives, also known as oligosaccharides, and in particular, to a method of identifying glycan structures using experimentally determined mass spectrometer data by correlating fragmentation patterns of glycan fragments with a theoretical or experimental database of glycan fragments.

Background of the Invention

It is known that it is possible to predict glycan structures from mass spectrometric fragmentation data using manual interpretive methods. These involve interpretation of the fragmentation of the sugar structures by assigning the loss of specific fragment masses produced by tandem mass spectroscopy. The unique properties of glycans, such as the possibility of numerous branch sites on each monosaccharide, as well as the isomers and anomers that exist, result in complex fragmentation spectra in which the fragments observed, result from both differential cleavage of the different linkages as well as internal cross-ring cleavages. The problem with such manual methods is that they are slow and time consuming. In order to enable high throughput characterisation of glycans by mass spectrometry it is necessary to provide a system which provides automatic comprehensive and rapid identification of glycan structures, and at the same time allow a non-biased interpretation of mass spectra, based on the interpreters knowledge.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed in Australia before the priority date of each claim of this application.

Summary of the Invention

In a first broad aspect of the present invention, there is provided a method of inferring the structure or sub-structure of a glycan or glycan derivatives by matching experimentally determined data, such as mass spectrometric fragmentation data to a data resource of theoretical peak masses.

In order to obtain theoretical peak masses for a Glycan structure, an algorithm is needed to generate fragments for the sets of n-cleavage points for a structure where n represents the number of cleavages needed to generate the fragment.

It is preferred that the method used for generating fragments is based on a combinatorial/permutation method.

Typically the method includes the steps of edge selection and cleavage assignment.

The data resource may comprise a theoretical database of fragments derived from a database of known glycan structures, a set of fragmentation rules, and/or a database of fragmentation structures which might be empirically determined.

A theoretical database of all possible fragments can be predicted from a database of known or theoretical glycan structures using software algorithms. Experimental fragmentation data can be collected by mass spectrometric techniques. That data can be used to generate a series of fragmentation rules which may be embodied in software algorithms to theoretically fragment the database of known glycan structures.

First, the experimentally determined mass (determined, for example, by mass spectrometry) of the glycan which is to be characterised or correlated with the database, is compared with the masses of the known structures in the database and those structures whose mass does not correlate with the experimentally determined mass within predetermined limits, are discounted to leave a reduced number of potential glycan matches.

In order to characterise oligosaccharides there are two criteria that need to be fulfilled; the sequence has to be identified and the linkage configuration and position has to be determined. Information about either of the two will provide valuable data. In a broad sense mass spectrometry will be able to predict the sequence information while linkage information will be more difficult to obtain with that technique. Cross ring cleavages or specific cleavages will have the potential to enable linkage position to be determined, while the linkage anomery is the parameter that is most difficult to obtain. On a computational basis, sequencing will be able to generate in silico glycosidic and cross ring fragments solely on a mathematical basis, but information about linkage anomery can not be included. The characteristic in a fragmentation spectra that has the potential of including some information about anomery is peak

intensity. This is purely since it could be envisaged that different anomeric configurations may undergo fragmentation rearrangements by different kinetics. Fragment intensities will of course also depend on other parameters. Scoring methods will be designed to do the following:

1. Provide a quality scoring based on the sequence allowing judgement of whether the sequence at least is correct.

2. Provide a ranking between oligosaccharides based on sequence (glycosidic cleavages)

3. Provide a ranking between oligosaccharides based on linkage position (cross ring cleavages)

4. Provide a ranking between oligosaccharides based on other cleavage types including generic n-cleavages and other special cleavage types where a special cleavage is a cleavage that produces a fragment that is specific to that structure which may include the loss of water, for example.

The initial data set used for comparison with the experimentally determined mass may consist of only fragments that are the result of 1-cleavage fragmentation as well as 2-cleavages which are formed exclusively from glycosidic cleavage types. The glycosidic cleavage pattern is the parameter that contains information about oligosaccharide sequence. This limited set of fragments provides enough data for the primary sequence scoring method to work. The increase in data set size by adding more fragments is limited by refining the data set when required. This way, by restricting the types of fragments generated based upon the results of the scoring, it is possible to keep the data set size to a manageable size.

In order to increase the accuracy of the method without sacrificing speed, the data set against which the method is performed is refined. For the initial data set used in the method , not all of the structures returned will be valid candidate structures for the spectrum. In order to exploit this, a more detailed method can be performed against the more likely structures out of the current result set. Extra fragments can be retrieved either from a slower secondary storage device, or generated on the fly for detailed queries. It is not necessary for the entire solution space for fragments to be available in every query, because the initial result that the query will provide is sequence correctness. By taking advantage of properties of the sugar structure fragmentation patterns, it is possible to target the data set for each query to contain only relevant data.

More specifically, the present invention involves the steps of obtaining experimentally observed fragmentation masses of a glycan, having an observed parent (un-fragmented) mass;

extracting a list of all glycans having the observed un-fragmented mass within a preset experimental error from a database comprising a list of all possible glycan fragments and their unfragmented molecular mass;

theoretically fragmenting the glycans in the list initially with a small subset of possible fragmentations, most preferably fragmentations resulting from a single cleavage or double cleavages formed exclusively from glycosidic bond cleavages; and

using a scoring method to rank each structure from the list to assess their likelihood of matching the experimentally fragmented glycan.

If it is not possible to determine a match from the limited set of fragments a second iteration may be carried out by adding more fragments such as triple cleavages to the data set.

Thus by restricting the types of fragment generated based on the results of the scoring it is possible to keep the data set to a manageable size.

The scoring method may be a quality scoring method most preferably utilising a grouping algorithm.

The fragment generation process will preferably omit redundant fragments and, when known, chemically impossible fragmentations to reduce the amount of fragments and data to be processed to make the method more efficient.

Thus, the present invention provides a method of identifying glycan structures from experimentally determined mass spectrometry data. In particular, automatic comprehensive and rapid identification of any glycan structure may be achieved using this technique from mass spectrometry fragmentation spectrum.

The identification of glycan differences offers indicators for recognition of glycosylation differences which for example can occur on proteins, lipids or proteoglycans. These variants have been linked to disease, cell differentiation, cell communications, immunological recognition and other significant characteristics.

Brief Description of the Drawings

Specific embodiments of the present invention will now be described, by way of example only, and with reference to the accompanying drawings in which:

Figure 1 illustrates the depth of edges in a glycan structure S where: depth (E_1) < depth (E_2) < depth (E_3) = depth (E_4)

Figure 2 illustrates disjoint fragments where edges E_3 and E_4 have been cut.

Figure 3 illustrates a non-disjoint double non-reducing end fragment;

Figures 4a to 4g schematically illustrate the method of glycan mass fingerprinting of the present invention;

Figure 5a is a graph showing a spectrum of peak masses of an experimentally fragmented oligosaccharide illustrating the fragments assigned to peaks in the spectrum;

Figure 6 shows an oligosaccharide structure of Example 1;

Figure 7 is a graph showing the spectra of the oligosaccharide structure of Figure 6;

Figures 8a to 8c are parts of a table giving the score, missed intensities and grouping score for a number of oligosaccharide structures which potentially match the oligosaccharide structure of Figure 6;

Figure 9 shows an oligosaccharide structure of Example 2;

Figure 10 is a graph showing the spectra of the oligosaccharide structure of Figure 9; and

Figures 11 shows a table giving the score, missed intensities and grouping score for a number of oligosaccharide structures which potentially match the oligosaccharide structure of Figure 9.

Detailed Description of a Preferred Embodiment

In the following specific description the following terminology applies.

Structure – An oligosaccharide consisting of monosaccharides connected by glycosidic bonds.

Peak – A peak in an MS/MS spectrum. This peak has a mass to charge (m/z) and relative intensity (relative to the largest peak in the spectrum).

Cleavage – A carbon bond in the structure is broken. The cleavage types are categorised into three types; glycosidic, cross-ring and special.

Single cleavage – A cleavage event that involves only a single glycosidic, cross-ring or special cleavage event.

Multiple cleavage – A cleavage event that involves more than one cleavage event. Can be described as n-cleavage events, ie. 1-cleavage, 2-cleavage etc.

Glycosidic cleavage – A cleavage involving the breakage of the glycosidic bond.

Cross-ring cleavage – A cleavage involving the breaking two of the carbon-carbon bonds in one of the carbon rings of a saccharide.

Special cleavage – A cleavage which is diagnostically significant, but does not directly fall into the glycosidic or cross-ring categories.

Fragment – The result of a single or multiple cleavage event.

Reducing end fragment – A fragment which contains the reducing end of the structure.

Non-reducing end fragment – A fragment which does not contain the reducing end of the structure.

In the present invention, a database of the theoretical peaks masses for all possible glycan fragments along with their unfragmented molecular parent mass, is produced by collating the set of theoretical fragments for an entire database of identified and characterised glycan structures. A number of suitable databases are available. For example, GlycoSuiteDB available at "www.glycosuite.com" provides a database of identified and characterised glycan structures as does the database "Glycominds". The theoretical fragmentation is carried out using a method set in more detail below.

In a refinement of the invention in order to match against and identify novel glycan structures, which are not already disclosed in existing databases, it is equally feasible to construct a theoretical database of all possible fragmentations of the much larger set of theoretically possible glycans. It is envisaged that this much larger database will be used for a second path search in which a glycan's fragment masses do not satisfactorily match to any known glycan fragment fingerprint.

The method used for generating fragments is based on a combinatorial/permutation method. The method can be broken into two stages namely edge selection and cleavage assignment.

In order to obtain theoretical peak masses for a Glycan structure, an algorithm is needed to generate sets of fragments for the full sets of n-cleavages for a structure.

Dealing with edge selection first, a structure S is composed of m monosaccharides with m-1 glycosidic bonds existing between monosaccharides. In order to generate a full set of fragments for n-cleavages, we need to consider the breakage of bonds at n positions (where n <= m-1). There exists $C^{m-1}{}_n$ combinations of glycosidic cleavage points (edges) for a n-cleavage fragmentation. In order to minimise size complexity an iterative method is used to generate all combinations of edges. E is the k-subset of the edges found in S. k can be any number up to (m-1).

For example the 2-subset is a set of all combinations of edges where two edges are combined. For the example shown in Figure 1 there are four edges E_1, E_2, E_3, and E_4. For a double cleavage, k=2 and the k subset comprises all possible combinations of E_1, E_2, E_3, and E_4, two at a time namely (E_1, E_2), (E_1, E_3), (E_1, E_4), (E_2, E_3), (E_2, E_4), and (E_3, E_4).

The edges within each k-subset are then sorted according to depth, which produces an edge vector. Edges that involve monosaccharides closer to the reducing end, are sorted with a higher rank than edges occurring at a greater depth. For example, with reference to Figure 1 illustrates the depth of edges in a glycan structure S where: depth (E_1) < depth (E_2) < depth (E_3) = depth (E_4). Edges E_1 and E_2 are selected from that Figure being the best two edges. The k-subset of edges is (E_2,E_1) and once sorted, the edge vector will be (E_1,E_2) since E_1 is closer to the reducing end of the structure which is conventionally drawn on the right of the structure and is the end in which the hydroxide on C-1 is not extended with additional monosaccharide units. The ordering of edges is crucial to ensuring the accurate generation of fragments, as it is possible to choose particular cleavages to assign to the edges so that a disjoint fragment is generated. Disjoint fragments are fragments which do not have any common monosaccharides. Thus with reference to Figure 2 if edges E_3 and E_4 are cut, two separate fragments are created.

Carbohydrate fragmentation patterns are discussed in the article "A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS

Spectra of Glycoconjugates" by Bruno Domon and Catherine E Costello published in Glycoconjugate J (1988) 5: 397-409, the entire contents of which are incorporated herein by reference. "Domon and Costello" notation is the accepted norm for labelling glycan fragment ions and is used herein.

Reducing end fragments may only be the result of particular types of cleavages. For 1-cleavages, these are the Y, Z, X and certain special cleavage types. For n-cleavages, reducing end fragments only occur where there are no B,C or A cleavages amongst the set of cleavages that occur. For example, reducing end fragments include Y , Z and Y/Z (Y and Z simultaneously) fragments. A B/Y fragment cannot be a reducing end fragment.

Non reducing end fragments can result from combinations of cleavage types that only include a single non reducing cleavage type. It is not possible to create a fragment from more than one non reducing cleavage type.

Calculating all the possible fragments is computationally intensive. Where two or more cleavages occur some of those 2-cleavages will produce fragments that will already have been accounted for in the 1-cleavages. For example the reducing end fragment produced when the two edges E_1 and E_4 are cut is also produced as a result of a 1-cleavage at E_1. The results of the E_4 cleavage are not used as E_4 did not reside on the reducing end fragment that was a result of the E_1 cleavage. Any fragments produced by a 2-cleavage, that are also produced by a 1-cleavage do not need to be calculated. Generally, when generating all n-cleavages, any fragments that could be produced by a m-cleavage (where $m < n$) are discarded.

For each combination of edges obtained in the edge selection step a fragment can be generated by applying a set of fragment types to it. Referring now to Figure 3 which shows a non-disjoint double non-reducing end fragment, consider a combination of edges formed from a 2-cleavage event consisting of Edge A and Edge B. At Edge A, the possible cleavage types that could have occurred are all reducing and non-reducing end cleavage types. At Edge B, only reducing end fragments could have occurred. Only reducing end cleavages occur at Edge B as it is not possible to have two non-reducing end cleavage types resulting in a non-disjoint fragment. A fragment of this type would in fact be identical to a single cleavage occurring at the edge B with the greatest depth.

To assign cleavages to fragments, we map the selection of cleavage types onto each element of E.

T = the Set of n element cleavage type permutations.

for $t \in T$, |t|=n

$\forall e \in E: \forall t \in T$: Fragment = (t,e) ie -(fragment type, position)

T is restricted so that each n-element permutation of cleavage types does not contain more than one non-reducing end fragment. Also, to avoid disjoint fragments occurring, the structure is checked to ensure that the structure can support the fragment. Basic checking occurs to invalidate any reducing end fragments where for a reducing end cleavage type assigned to a cleavage point, a traversal to the reducing end of the structure does not traverse any other cleavage points. Non-reducing end fragments are marked as invalid if for any of the reducing-end cleave points a traversal to the reducing end does not pass a B cleavage point. Checking occurs by starting at the cleavage point occurring at the least depth (closest to the reducing end), traversing the structure towards the reducing end, and marking any monosaccharide that is traversed over. This is repeated for the other cleavage points in the fragment. Any fragment which causes the loss of branches containing marked monosaccharides due to an A cleavage type is discarded.

Once the assignment of cleavage types to cleavage points has been verified, a virtual fragmentation occurs of the structure. This process involves removing branches from the virtual representation of the structure so that it will represent the structure of the fragment. Once the virtual fragment has been generated the mass can be obtained by looking up the masses of the remaining monosaccharides, as well as any mass losses of fragmentation types. An identifier for this fragment is created based upon the Domon + Costello notation and assigned to the fragment.

The generation of fragments is a combinatorially difficult problem. As the number of fragments dramatically increases as the number of allowed cleavages increases, it is not feasible to generate all fragments a-priori. The method of the present invention is initially performed against a smaller subset of theoretical fragments which are stored in a database. Typically the fragments for 1-cleavages, and 2-cleavages from exclusively glycosidic cleavages will initially be used.

The basic process referred to hereinafter as Glycan Mass Fingerprinting or GMF, involved in characterising an oligosaccharide is a s follows. A user will supply a spectrum, which consists of tuples of m/z and intensity values. Each tuple is called a peak. The peak mass is converted into a true mass_by

adjusting for charge state and adduct, and then compared against the set of theoretical fragments to find any fragments which have a mass within the tolerance range of the peak's true mass. The fragments are then collated according to the parent structure and scored. Based on the scoring, the data set size is further increased by adding more fragment types, and the process is performed again. The process can only be repeated until the experimental data set is exhausted of the required information, ie no unique fragments can be found that distinguish particular oligosaccharide candidates. In order to improve efficiency, only a portion of the spectrum may need to be used, or the process may only be performed against fragments which are the result of certain structures being fragmented. A structure which has at least one fragment which matches with a peak true mass will have a set of fragments associated with it. This fragment set is the set of fragments derived from the structure which have matched with the spectrum peak true masses.

The initial data set used for GMF consists of only fragments that are the result of 1-cleavage fragmentation as well as 2-cleavages which are formed exclusively from glycosidic cleavage types. This limited set of fragments provides enough data for the primary sequence scoring method to work. The increase in data set size by adding more fragments is limited by refining the data set when required. This way, by restricting the types of fragments generated based upon the results of the scoring, it is possible to keep the data set size to a manageable size.

There are two scoring methods used, one to determine the sequence quality of match of a candidate structure, and another to rank the candidate structures relative to each other. The family of algorithms for each scoring type are defined as quality and relative scoring methods respectively. Based on the combination of these two scoring methods, it is possible to determine the likelihood of a result structure being the one defined by the input spectrum, in regards of sequence or linkage information or both.

The quality score for a result encapsulates how well the fragments matched for a sequence define that sequence. For example, a result structure that matches only a single small fragment will be a low quality result, whilst a structure which has many fragments matched which are distributed over the entire structure will have a high quality score. One such quality scoring algorithm is a grouping algorithm.

Grouping scoring derives the cleavage points from the fragment types, and obtains a number which represents how well the structure is characterised by the set of fragments associated with it. The best fragments used to characterise a structure are those resulting from 1-cleavages. If there are m - 1 unique cleavage points found in a glycan structure's associated 1-cleavage fragments for a glycan having m monosaccharides, then there is enough evidence in the fragments that the sequence of the structure is valid.

Fragments resulting from 2-cleavages do not necessarily indicate the presence of a specific cleavage point in a structure. 1-cleavages are special as the presence of a fragment is enough evidence to prove that a fragment occurred at the cleavage point. 2-cleavages can be considered as a fragmentation of a fragmentation. One of the cleavage points in a 2-cleavage can be used as evidence if the other cleavage point has evidence supporting it's existence. In other words, the 2-cleavage must have an overlap with another 1-cleavage, or 2-cleavages where one of its cleavages have been assigned, for it to contain an equal amount of information. For this reason, 2-cleavages are not weighted as importantly as 1-cleavages. Any scoring method that examines cleavage points should be able to encapsulate this information. One possible algorithm involves a process of trying to fulfil each cleavage point in the original structure with a matched fragment. Whenever possible the grouping scoring algorithm will try to use a single cleavage fragment to fulfil the cleavage point. If the cleavage point cannot be fulfilled by a 1-cleavage fragment, it will use a 2-cleavage fragment. The actual score assigned is derived using:

Equation 1

$$\text{Score} = (a - 0.25b) / (m - 1)$$

where a is the number of cleavage points assigned to 1-cleavage events, and b is the number of cleavage points assigned to 2-cleavage fragments. A structure whose cleavage points are strongly supported by it's fragments is assigned a score closer to 1. This method can be extended to handle generic n-cleavages where n is greater than 1, by extending the formula to appropriately weight the importance of the cleavages and further subtracting those from a.

It should be noted that the above is only one simple type of scoring equation and that other equations could be used to perform the same function encapsulating the information from both single and double cleavages.

Relative scoring methods will allow for differentiation of results which have the same quality score. One method which can be used is a matched intensity scoring method. Matched intensity can also be further refined into matched sequence (only glycosidic cleavages) intensity and linkage information (cross ring, special cleavages with or without concomitant glycosidic cleavages) intensity.

Matched intensities obtains the sum of intensities of all peaks which have matched with at least one fragment within a fragment subset (eg glycosidic, cross ring, or both together). A peak matching with at least one fragment suggests that there is a possible fragmentation that can support this peak mass. Structures which are more correct will have a greater number of spectrum peaks matching with any fragments. The matched intensity score is particularly useful for distinguishing between isomers of structures, which may otherwise have an identical grouping score. The matched intensity score will determine the quantity of diagnostic fragments that have matched, and a difference in score suggests a difference in matched fragments. For ease of reporting, the matched intensity score can be converted into a missed intensity score, which is simply the sum of total intensities in the spectrum minus the matched intensity score.

In order to increase the accuracy of the GMF process without sacrificing speed, the data set against which GMF is performed is refined. For the initial data set used in GMF, not all of the structures returned will be valid candidate structures for the spectrum, as they may not have the right sequence. In order to exploit this, a more detailed GMF can be performed against the more likely structures out of the current result set. Extra fragments can be retrieved either from a slower secondary storage device, or generated on the fly for detailed GMF queries. It is not necessary for the entire GMF solution space for fragments to be available in every GMF query. By taking advantage of properties of the sugar structure fragmentation patterns, it is possible to target the data set for each GMF query to contain only relevant data.

As the data sets become more refined, and the possible solution set more relevant, the matched intensity score will increase. Initial data sets will contain generic fragments, and will not match more exotic fragments which may occur. However, these exotic fragments may not necessarily be useful in determining the correct result out of a large result set. For example, the intensity of the peak matching the fragment may be very low, or the fragment

occurs in many of the structures. As the result set is reduced in size the importance of these fragments increases, and they play a very important role in the selection of the most probable candidate structure.

Figures 4a to 4g schematically illustrate the general method used to perform glycan mass fingerprinting (gmf). Individual oligosaccharides could be submitted to gmf after mass spectrometry under conditions producing fragment ions for example by tandem mass spectrometry, or in source fragmentation, or alternatively oligosaccharide mixures could be separated into individual components with separating methods hyphenated with mass spectrometry. This includes techniques such as hplc and capillary electrophoresis. Various ionisation methods and conditions could be used. Multiple stages of mass spectrometry could also be used, where further fragmentation of fragment ions is required. Figure 4a shows a parent molecular ion mass from an MS spectrum of the unknown glycan which requires further investigation.

The glycan database (Figure 4b) consists of a set of glycan structures and their associated masses. The database can be in simple table form, or can be in a relational form to exploit other information that may be associated with glycan structures such as biological source information.

Figure 4c shows a spectra obtained from fragmentation of the unknown glycan. The spectra is plotted to show the relationships between mass/charge ratio and intensity per peak.

Next candidate structures are theoretically fragmented (refer to Figure 4d) in order to allow matching between the peak masses and fragment masses using the processes described above. The database is initially queried to restrict possible structures to only those which have the same mass as the molecular ion within preset tolerances.

The fragments may be pre-generated and stored in a database, or generated as required or both.

The spectra is iteratively matched and scored against the set of fragments (Figure 4e). At the end of the matching and scoring stage, a ranked set of results is obtained (Figure 4f). The highest ranked structure is a structure from the database which is most likely to be the correct structure.

Figure 5 shows a graph of peaks from fragmentation of a glycan structure 10. Peak m/z 689.9 has been matched with two different fragments having the same mass. Further information is required to determine whether both the fragments that have matched, or a single one is the correct fragment.

Example 1

Figures 6 to 8 illustrate a first example. The oligosaccharide structure which is empirically fragmented is shown in Figure 6. Figure 7 shows its m/z spectra. Figures 8a to 8c show a table of results illustrating how the method can distinguish between two isoforms of structure when the grouping score is the same by comparing the sum of the missed intensities with the first structure being the correct structure and having a lower total sum of missed intensities despite both structures having the same score of 0.8 as determined by equation 1.

Example 2

Figures 9 to 11 illustrate a second example. The oligosaccharide structure which is empirically fragmented is shown in Figure 9. Figure 10 shows its m/z spectra. The first result on this table is correct as it has both a perfect grouping score and the lowest number of missed intensities.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

Dated this eleventh day of June 2003

Proteome Systems Intellectual
Property Pty Ltd
Patent Attorneys for the Applicant:

F B RICE & CO

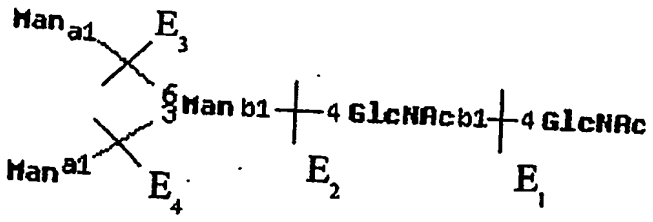


Figure 1

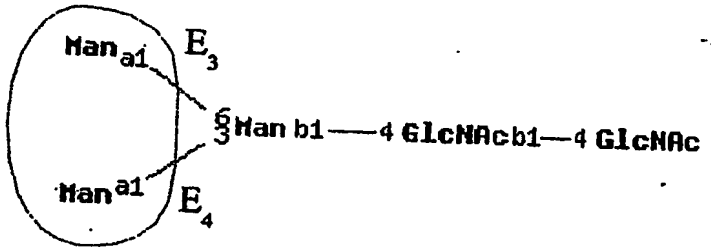


Figure 2

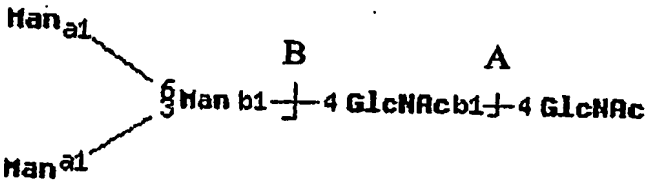


Figure 3

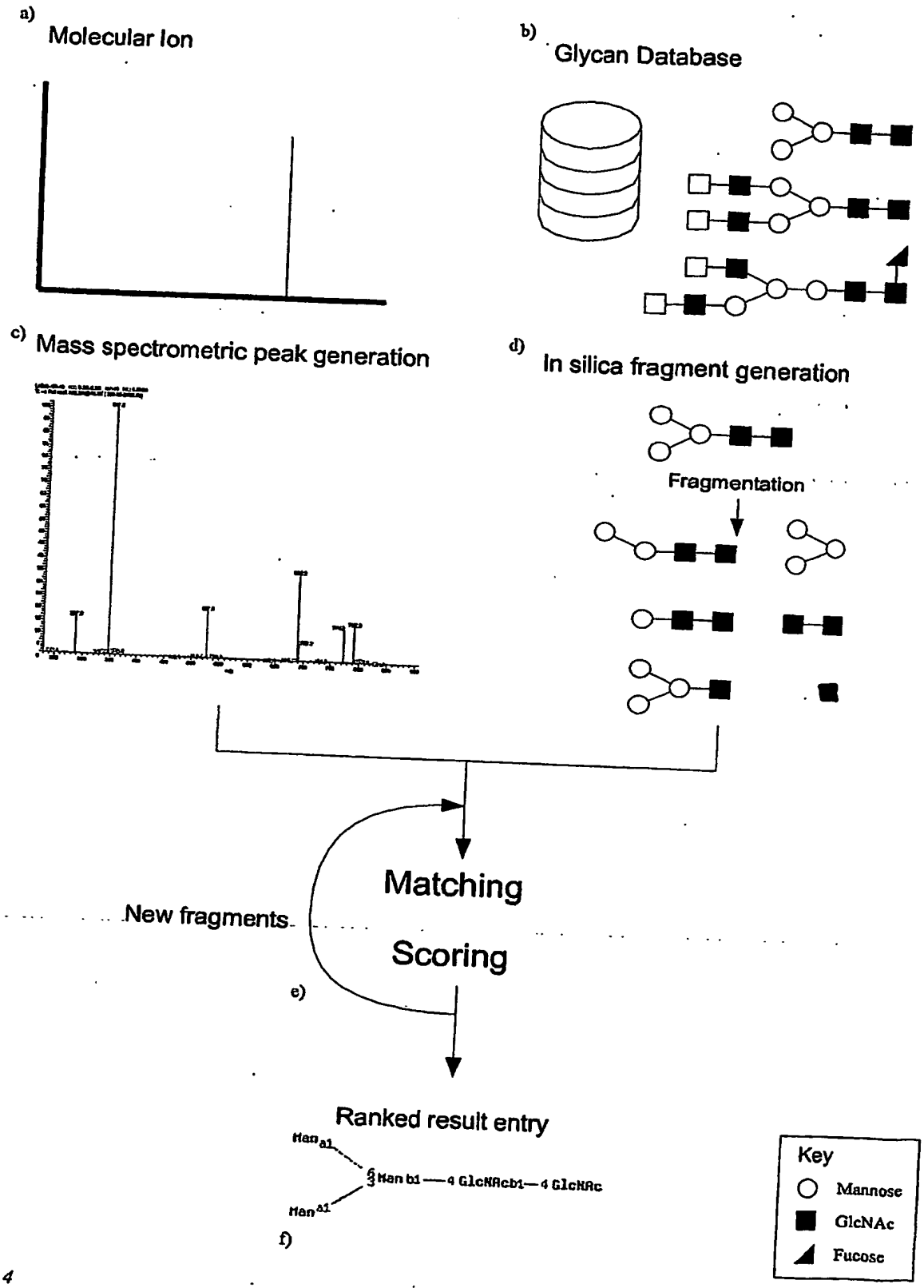

a)
Molecular Ion
b)
Glycan Database
c)
Mass spectrometric peak generation
d)
In silica fragment generation
Fragmentation
Matching
New fragments
Scoring
e)
Ranked result entry
Man α1
Man α1
Man b1 — 4 GlcNAcb1 — 4 GlcNAc
f)
Key
Mannose
GlcNAc
Fucose

Figure 4

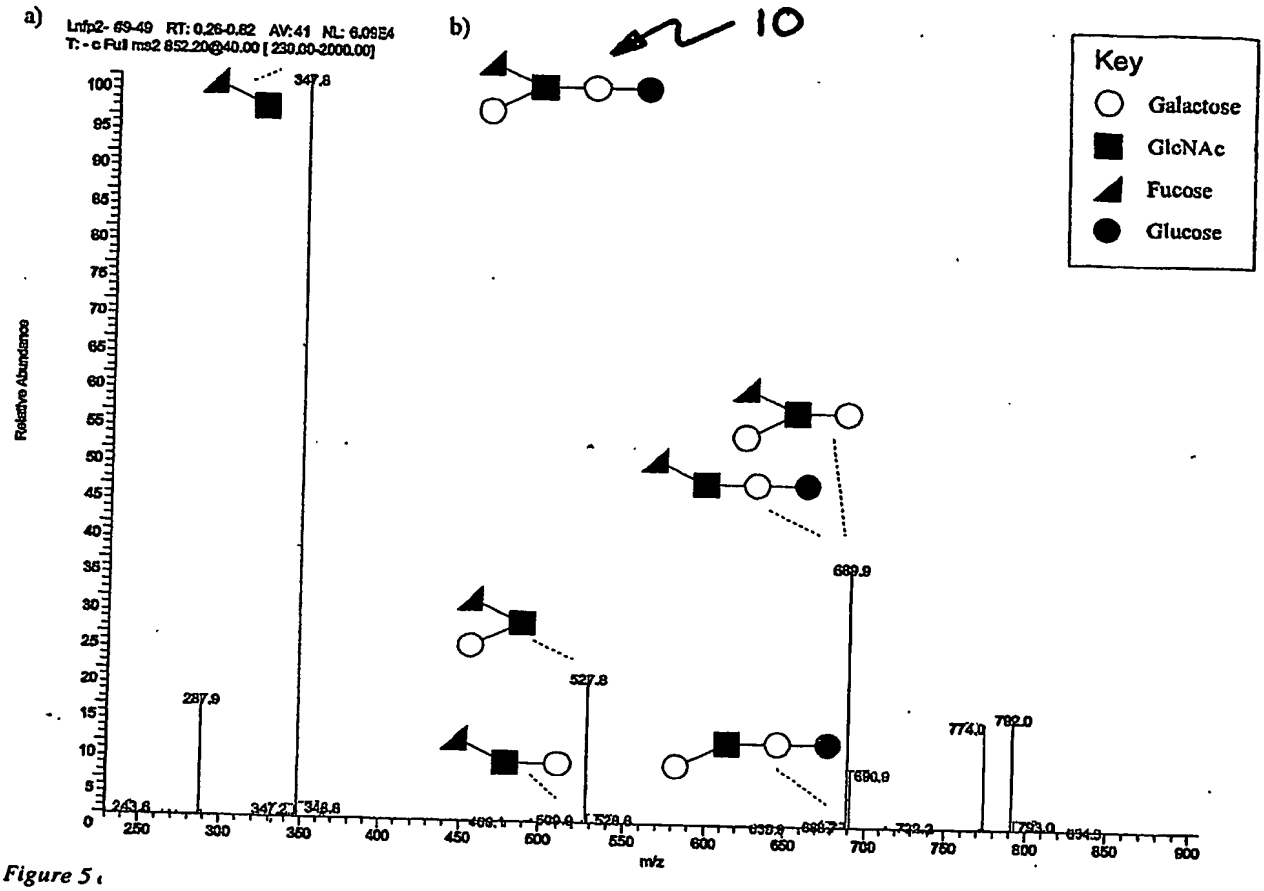

a)
Lnfp2- 69-49 RT: 0.26-0.82 AV:41 NL: 6.09E4
T: - c Full ms2 852.20@40.00 [230.00-2000.00]
Relative Abundance
m/z
b)
10
Key
Galactose
GlcNAc
Fucose
Glucose
347.8
287.9
527.8
689.9
650.9
774.0
792.0

Figure 5

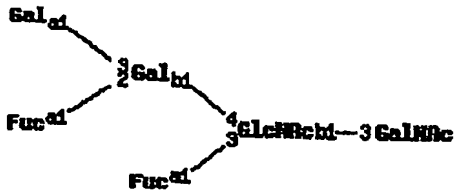


Figure 6 Oligosaccharide structure

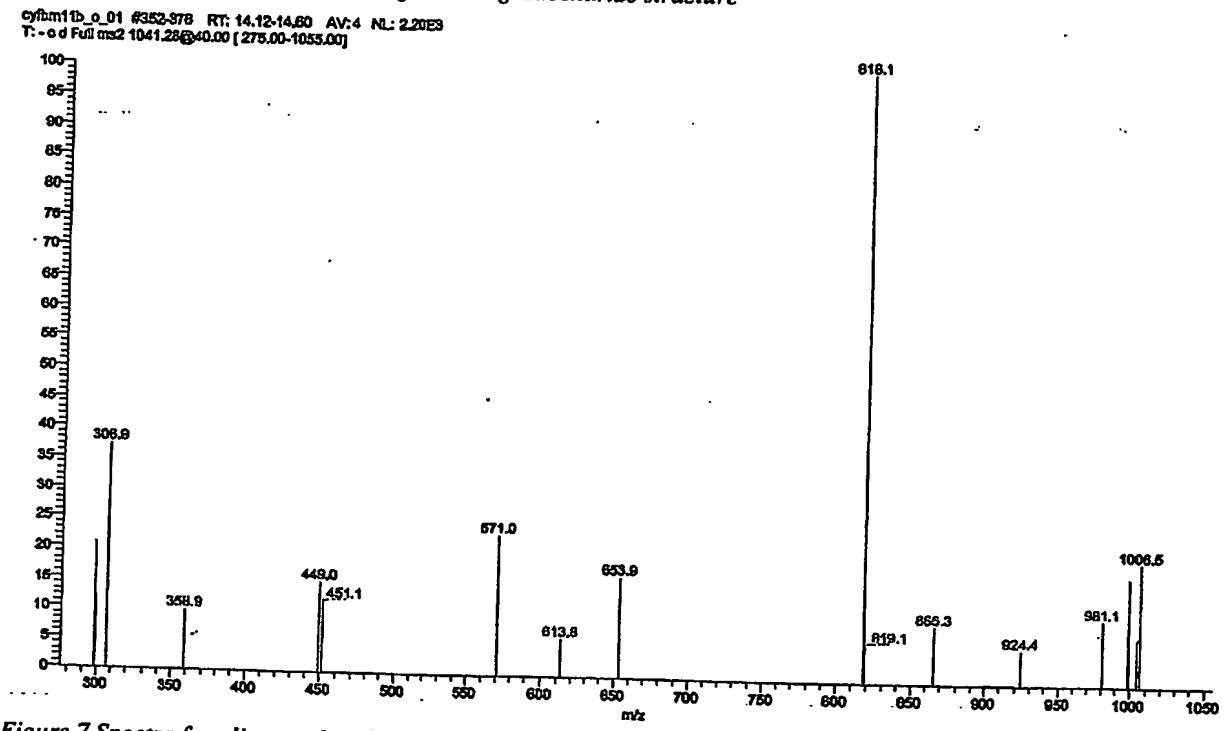


Figure 7 Spectra for oligosaccharide shown in Figure 6

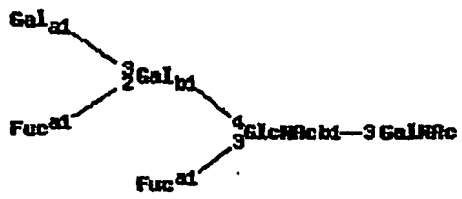
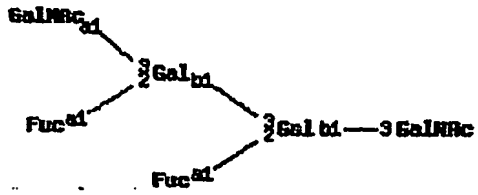
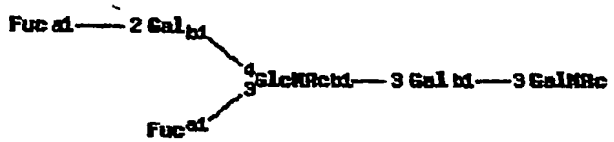
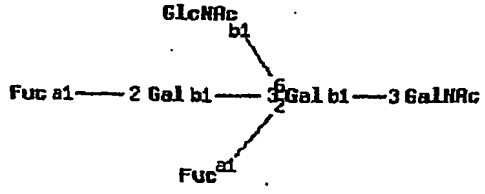
Score	Missed intensities	Grouping score	Structure
.8	123.6	5 out of 5 using 4 double fragments	
.8	156.32	5 out of 5 using 4 double fragments	
.7	141.82	4 out of 5 using 2 double fragments	
.7	147.65	4 out of 5 using 2 double fragments	
.7	162.15	4 out of 5 using 2 double fragments	

Fig 8a

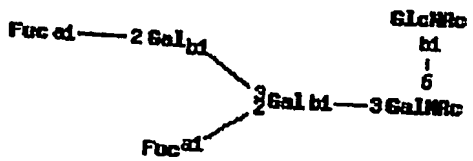
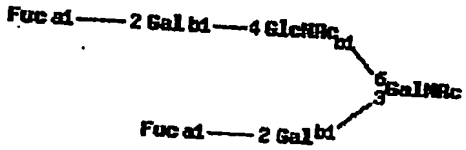
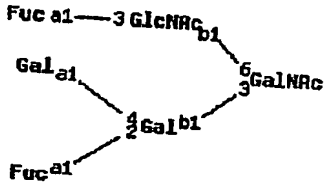
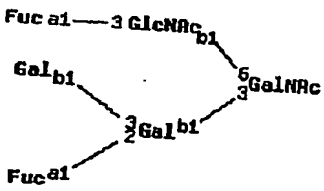
Score	Missed intensities	Grouping score	Structure
.65	266.98	4 out of 5 using 3 double fragments	
.45	286.4	3 out of 5 using 3 double fragments	
.4	278.51	2 out of 5	
.35	232.08	2 out of 5 with 1 double fragment counted	
.35	232.08	2 out of 5 using 1 double fragment	

Fig 8b

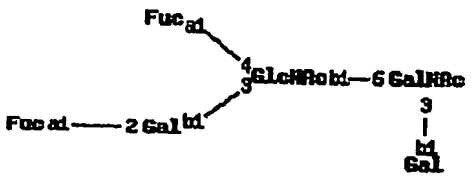
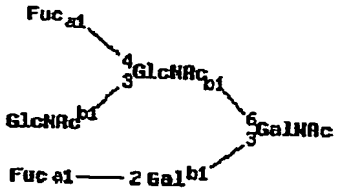
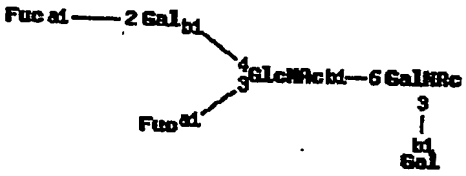
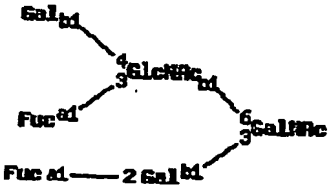
Score	Missed intensities	Grouping score	Structure
.35	235.24	2 out of 5 using 1 double fragment	
.35	235.57	2 out of 5 using 1 double fragment	
.35	241.07	2 out of 5 using 1 double fragment	
.35	255.57	2 out of 5 using 1 double fragment	

Fig. 8c

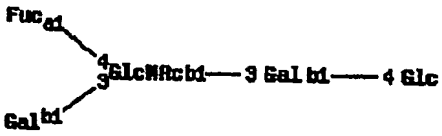


Figure 9

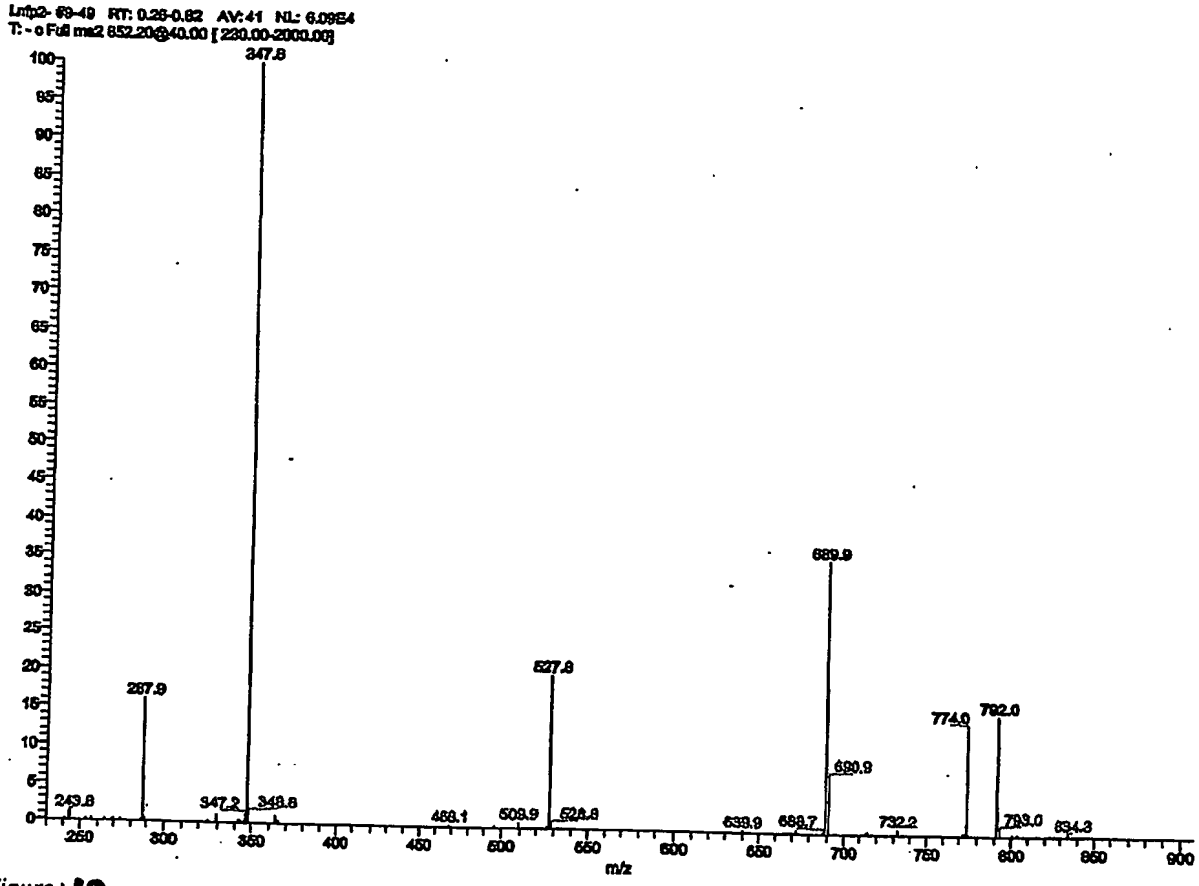


Figure : 10

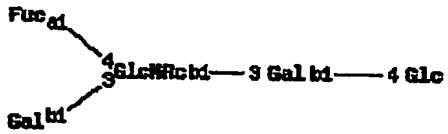
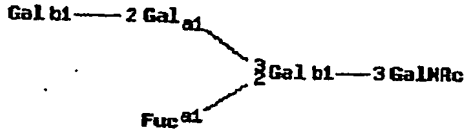
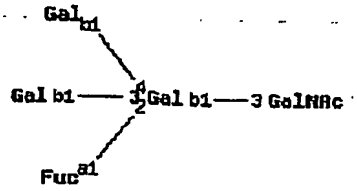
Score	Missed intensities	Grouping score	Structure
1	29.91	4 out of 4	
1	131.43	4 out of 4	Fuc a1——2 Gal b1——4 GlcNAcb1——3 Gal b1——4 Glc
1	131.43	4 out of 4	Fuc a1——2 Gal b1——3 GlcNAcb1——3 Gal b1——4 Glc
1	131.43	4 out of 4	Fuc a1——2 Gal b1——3 GlcNAcb1——3 Gal b1——4 Glc
.9375	124.3	4 out of 4 using 1 double fragment	
.75	124.87	3 out of 4	
.375	190.01	2 out of 4 using 2 double fragments	GalNAcb1——4 GlcA b1——3 Gal b1——3 Gal b1——4 Xyl

Fig 11